Gender differences in humor-related traits, humor appreciation, production, comprehension, (neural) responses, use, and correlates: A systematic review

Hofmann, Jennifer ; Platt, Tracey ; Lau, Chloe ; Torres-Marín, Jorge

Abstract: All available peer-reviewed literature on humor and gender differences (1977–2018) was screened and evaluated according to a priori defined QUALSYST criteria. The 77 papers surpassing a conservative quality criterion generated seven emergent themes around humor and gender differences. In short, men score higher in the aggressive humor style (M > F), while no other gender differences were consistently reported in humor-related traits (M = F). In the prediction of negative outcomes (stress, loneliness, depression), differential effects for humor in both genders are reported, but not consistently (M F). Gender differences exist for the appreciation of sexual humor (M > F), even in mixed target stimuli, and hostile humor (both genders appreciate opposite gender target stimuli more). Gender differences are absent in nonsense and neutral humor (M = F). For humor production, three samples showed no gender differences (M = F), while three samples suggested men are funnier (M > F) and one that women are funnier (M < F). No studies reporting differences in humor comprehension were identified (M = F). For humor use and communication, gender differences were found across methods (M F), yet, they depend on the context (e.g., workplace) and may thus resemble gender roles rather than "natural differences". Moreover, few studies provide hard data on actual humor use and communication in different domains. When exposed to humor stimuli, different neural responses of men and women in prefrontal cortex activations (or selected parts) were found (M F). Also, self-report data suggest that both genders value a sense of humor in their partner (M = F), yet women typically value the humor production abilities more than humor receptivity, while for men, the woman's receptivity of their own humor is more important than a woman's humor production abilities, in line with gender stereotypes (M F). To conclude, much progress has been achieved in the past 15 years to overcome methodological flaws in early works on humor and gender differences. Importantly, attention should be paid to disentangling actual gender differences from gender role expectations and gender stereotypes. Methodologically, designs need to be checked for potential bias (i.e. self-reports may accentuate roles and stereotypes) and more hard data is needed to substantiate claims from self-report studies.

DOI: https://doi.org/10.1007/s12144-020-00724-1

DOI: https://doi.org/10.1007/s12144-020-00724-1

Abstract

All available peer-reviewed literature on humor and gender differences (1977-2018) was screened and evaluated according to a priori defined QUALSYST criteria. The 77 papers surpassing a conservative quality criterion generated seven emergent themes around humor and gender differences. In short, men score higher in the aggressive humor style (M>F), while no other gender differences were consistently reported in humor-related traits (M=F). In the prediction of negative outcomes (stress, loneliness, depression), differential effects for humor in both genders are reported, but not consistently (M≠F). Gender differences exist for the appreciation of sexual humor (M>F), even in mixed target stimuli, and hostile humor (both genders appreciate opposite gender target stimuli more). Gender differences are absent in nonsense and neutral humor (M=F). For humor production, three samples showed no gender differences (M=F), while three samples suggested men are funnier (M>F) and one that women are funnier (M<F). No studies reporting differences in humor comprehension were identified (M=F). For humor use and communication, gender differences were found across methods (M≠F), yet, they depend on the context (e.g., workplace) and may thus resemble gender roles rather than "natural differences". Moreover, few studies provide hard data on actual humor use and communication in different domains. When exposed to humor stimuli, different neural responses of men and women in prefrontal cortex activations (or selected parts) were found (M≠F). Also, self-report data suggest that both genders value a sense of humor in their partner (M=F), yet women typically value the humor production abilities more than humor receptivity, while for men, the woman's receptivity of their own humor is more important than a woman's humor production abilities, in line with gender stereotypes (M≠F). To conclude, much progress has been achieved in the past 15 years to overcome methodological flaws in early works on humor and gender differences. Importantly, attention should be paid to disentangling actual gender differences from gender role expectations and

gender stereotypes. Methodologically, designs need to be checked for potential bias (i.e. self-reports may accentuate roles and stereotypes) and more hard data is needed to substantiate claims from self-report studies.

Keywords: Gender Differences, Humor, Sex, Gender Roles, Individual Differences

Gender Differences in Humor-Related Traits, Humor Appreciation, Production, Comprehension, (Neural) Responses, Use, and Correlates: A Systematic Review

Individual differences exist in all aspects of humor: individuals differ in habitual aspects relating to humor (i.e., their "sense of humor"), as well as aspects relating to humor appreciation, comprehension, production and communication (for overviews see Martin, 2010; Ruch, 1998, 2008). In the comprehension of humor, or its appreciation, the same joke will not be understood or found funny by all people. When communicating humor, some individuals will rely on recalling other people's jokes, while others will create new, or embellish, real events to create funny situations. Some individuals will initiate humorous moments and situations, whereas others, while enjoying them, may not instigate them.

Across several disciplines, it was suggested that humor differs between genders. Yet, there are inconsistencies, even disagreement, about the magnitude of those effects, their exact nature, and their universality versus cultural specificity, and role dependency. Moreover, methodological challenges, for example a lack of nomenclature for "sense of humor" across time, places and cultures, or the choice of humorous materials to investigate humor appreciation, make it difficult to compare results across studies. Thus, the results from single independent articles do not allow for generalizable findings.

Lampert and Ervin-Tripp (1998) brought together and summarized the state of the art of the literature on gender and the sense of humor twenty years ago. They utilized existent literature to explore and summarize the gender difference findings. Structuring their results chronologically, they argued that "theories centered on socialization, social status and dispositional attitudes" (p. 237) accounted for findings regarding gender difference in humor appreciation in the earliest literature (pre-1970's). These early results suggest men use more humor than women, men appreciate humor more than women, especially aggressive and sexual humor, and both genders display more humor directed at female than male targets. Summarizing the quasi-experimental post 1970's findings (up to 1997) and structuring them

along the utilized study designs and methods, the authors suggest the "emergence of a trend among women and girls toward a diminished appreciation for female-targeted and increased acceptance of male-targeted humor" (pp. 244). Socioeconomic status, ethnicity, and age were also acknowledged as themes. No gender differences were found for personality related aspects of humor and similar findings were recorded for self-reported joke preferences but noted that "men and women enjoy jokes about sexual relationships and sometimes tell similar jokes, but their jokes do not serve the same psychological or interpersonal functions" (pp. 249). Lampert and Ervin-Tipp (1998) also introduced the notion that gender differences might be more prone in "lab settings" – yet less evident in "natural settings" (such as conversations). Concluding, the paper highlights the limitations to the range of settings and behaviors investigated, as well as methodological flaws (i.e., biased stimuli, utilizing self-reports to assess abilities, etc.). The outcome of which is the exaggeration of actual gender differences (see also Ergül, 2014).

From a pragmatics perspective and utilizing a collection of papers spanning 60-years relating to conversational humor, Kotthoff (2006) identified four themes from the literature on humor and gender: sexuality, status, aggressiveness, corporeality and social alignment that had gender differences within the humor practice. Kotthoff highlights a number of research gaps that remained in 2006 to adequately illuminate the differences in humor studies across the genders.

The current study aims to update the review of Lampert and Ervin-Tripp (1998) and work of Kotthoff (2006) utilizing a systematic review approach on the peer-reviewed literature between 1977 and 2018 to scrutinize the rigor of the papers and investigate whether new themes have emerged or stabilized over time. The focus is to identify, present and discuss all available and relevant literature that focuses on gender differences in adults in all aspects relating to humor, across disciplines and including quantitative, as well as qualitative approaches. While some aspects of humor, such as the sense of humor as a personality

characteristic, may be rooted in temperament (and biology to some extends), they are influenced by the interaction with the environment and social context. Moreover, some differences may be explicitly linked to certain gender roles and cultural specifics (such as cultural display rules, politeness norms, etc.), that do not relate to actual differences between men and women, but rather norms and roles that are carried out (to different extends). Where possible, this review also tries to distinguish such effects. Lastly, this review focuses on the very simplistic distinction between men and women. Following the descriptions of the American Psychological Association (2009), we use the term gender as "gender is cultural and is the term to use when referring to woman and men as social groups" (p. 71), as opposed to sex and sex differences, as "sex is biological" (p. 71) and the term should be used when the biological distinction between men and women is predominant.

Method

Framing Questions, Identification of Literature and Inclusion/Exclusion Criteria

This systematic review adhered to the five-step approach protocol outlined by Khan, Kunz, Kleijnen and Antes (2003): First, questions are framed that the review tries to answer. The framing questions for the current review were:

(1) In what way has humor been explored and investigated in relation to gender and gender bias through empirical literature? What themes emerge around the research conducted on humor and gender differences?

(2) What is the quality of the evidence around humor and gender differences?

(3) Is there a role in research methodology on the findings of gender differences?

Second, relevant publications were identified. The search strategy included a broad literature review on the following databases platforms: Web of Science (SCI-EXPANDED, SSCI, and A&HCI), EBSCO, Medline, Cinahl and PsycINFO databases between 1977 and the 1st of February 2018. All combined terms with the Boolean operator "and" in the title or

abstract were identified ("humour" OR "humor" OR "humourous" OR "humorous" OR laughter OR laughing OR laugh OR fun OR funny OR jokes OR joking OR comedy OR comedian OR wit) AND (gender OR gender bias OR bias OR gender differences OR sex OR sex differences OR gender roles OR sexist OR man OR woman OR female OR male OR men OR women). All citations were downloaded onto EndNote X8 and duplicates from all the databases were deleted using the "find duplicate" function. Duplicates that were not detected by the software (e.g., due to entry variation in some research databases) were deleted manually. Figure 1 depicts the process of literature identification.

Insert Figure 1 about here

As Figure 1 shows, the titles and abstracts of all unique citations were reviewed to determine which studies fit the review's pre-specified eligibility criteria. Articles included in the study either (1) identified their primary research question as humor and gender differences, as suggested by the title and abstract of the study, (2) the abstract suggested there were gender differences in the full manuscript worth noting (e.g., "significant gender differences emerged in the findings…"), and (3) the main research question was not to investigate gender differences, but a substantive section of the full text was dedicated to investigating humor and gender differences. The third criterion was added as many studies conducted secondary analyses on gender differences and had these studies been excluded, some valuable information would have been lost.

Inclusion criteria for this study are as follows:

(1) Empirical articles with primary data (utilizes sophisticated methodology, experimental designs, psychometrically sound assessment tools, and statistical analysis or systematic treatment of the data) published in an academic journal with a peer review process.

(2) Articles focused on male and female differences (including differences in dyadic relationships, such as romantic relationships, marriages) with respect to humor as a primary focus of the paper, or reported as key findings. Each article was evaluated individually, based on the aforementioned criteria.

(3) Papers with secondary analyses of gender differences must analyze male and female differences with respect to humor in a statistically sound method. A substantive portion (i.e., a separate section or paragraph) had to be dedicated to investigating differential gender differences. Each article was individually scrutinized to ensure it met the exclusion criteria as well.

Two authors then screened the title and abstracts of all other identified citations for potential eligibility. If eligibility criteria could not be determined based on the abstract and title alone, the full manuscript was searched for “male”, “female”, “men”, “women”, “boy”, “girl”, “gender”, and “sex” to identify information on gender differences. The exclusion criteria are shown in Figure 1. Articles were excluded if they are: (1) including a different definition of humor (i.e., aqueous humor of the eye), (2) not peer reviewed or peer review status unclear, (3) reviews, letters, commentaries, editorials, meeting or conference abstracts, (4) studies that focused on sexuality, rather than sex differences, (5) articles on laughter as a general nonverbal signal or laughter propensity without the link to humor or ridicule (these studies were excluded because although humor and laughter might overlap, humor occurs without laughter and there is laughter that is not related to humor, see Hofmann & Ruch, 2017), (6) unspecified laughter as the only outcome measure, (7) studies on children (as we wanted to focus on adults, i.e., individuals aged more than 18 years), (8) clinical studies on specific patient groups or case reports (9) if the entire study comprised only males or females in the sample (making a direct comparison between the genders impossible) or no specification of participant gender in the sample description, (10) analysis of humor and gender roles in fictional characters in films, television series, operas books, or specific

comedians, (11) including one item humor measures, (12) jokes as an operationalization of a different concept (e.g., sexual harassment) (13) decoding studies without the assessment of inter-individual differences, (14) studies with limited possibilities for generalization of the results (e.g., on gender roles in specific subgroups/populations, on particular cultural displays). There were no restrictions with respect to countries or journals. After this process, a divergence of 35 cases (1.67% of the articles) was detected between the two raters and solved upon discussion and consent, leading to a total of 158 eligible for the further assessment of the study quality (QUALSYST rating).

Assessing the Study Quality

The standard quality assessment was conducted to evaluate primary research papers (Kmet et al., 2004). Risk of bias was assessed using the quality assessment tool 'QUALSYST' from the "Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields" (Kmet et al., 2004). The QUALSYST assessment utilizes items evaluating specific criteria for research quality, including qualitative approaches (assessed with 11 items) and quantitative approaches (assessed with 14 items, of which three target intervention studies). Qualitative and quantitative papers were analyzed separately. The quality of research studies with quantitative data involved analyzing the representativeness and adequacy, sample size, the study design and methodology, measures used to assess humor, confounding variables, whether variance was reported, results reported in sufficient details, and conclusions supported by results (scale: yes = 2, partial = 1, no = 0; Kmet et al., 2004). The three items specifically targeting intervention studies (i.e., If interventional and random allocation was possible, was it described? If interventional and blinding of investigators was possible, was if reported? If interventional and blinding of subjects was possible, was it reported) were not applicable to most rated studies and thus not included in the score. The quality of research studies with qualitative data involved analyzing the following questions: (1) Question/objective clearly described? (2) Design evident and

appropriate to answer the study questions? (3) Context for the study is clear? (4) Connection to a theoretical framework/wider body of knowledge? (5) Sampling strategy described, relevant and justified? (6) Data collection methods clearly described and systematic? (7) Data analysis clearly described, complete, systematic? (8) Outcome and if applicable exposure measure well defined and robust to measurement/misclassification bias? (9) Use of verification procedure(s) to establish credibility of the study? (10) Conclusions supported by the results? (11) Reflexivity of the account?

Following the procedure by Kmet et al. (2004) a quality mean score was calculated across all applicable criteria items for each article (scores ranging from 0-1). Kmet et al. (2004) consequently suggest scores of >0.75 indicators of strong quality, 0.55-0.75 indicators of moderate quality, and <0.55 indicators of weak quality. For the current study, we chose the conservative cut-off of 0.75 as a criterion for study inclusion. A data extraction form was used to summarize evidence for each study that was evaluated, which included number of subjects (% female), study design, country the study was conducted, how humor was measured, and key findings related to gender differences. The content was used to create a narrative for emerging themes from the literature. The themes can help provide an overview for the major areas of research. Based on the quantity and quality of studies reviewed, conclusions to the proposed research questions were constructed. Formal meta-analysis was not possible due to differences between studies in the approaches used to measuring humor that were being compared. Instead, given that data pooling is inappropriate from measurement, statistical, and methodological heterogeneity, a narrative synthesis of the findings is provided. Tables illustrate key findings related to gender differences and study demographic (see Tables 1 to 7 in the Appendix).

Results

Quality Assessment

Quality assessment criteria were applied by two authors to each of the 158 identified articles (see Figure 1), the correlation of the total scores was r (158) = .744 p < .001. First, inclusion (cut-off of > 0.75) and exclusion rates were determined for both ratings separately, leading to an overall classification agreement of 78.47%, with 58 articles excluded by both raters (36.70% of the eligible articles) and 66 articles included by both raters (41.77% of the eligible articles). Moreover, 34 articles (21.53%) led to divergent decisions. To resolve the divergences, the decisions were compared to the mean rating across the two raters and the decision of inclusion guided by the mean rating (inclusion: ≥ .75; exclusion: < .75). With this procedure, 11 articles could be assigned to the inclusion category, as the mean rating suggested inclusion and additionally, the two ratings only slightly diverged (e.g., rater one having a mean rating of .73 and rater two having a rating of .82) or agreement was reached after discussion, and 23 were rejected. This led to a final set N = 77 included articles of, see Figure 1.

Theme Building

After screening the contents of the articles seven themes emerged: Differences in trait sense of humor and humor styles (N = 21), differences in humor appreciation (N = 24), differences in humor production and comprehension ability (N = 6), differences in neuroanatomical responses (N = 3), difference in humor use and communication (N = 11), differences in responses to humor (N = 9) and differences in correlates of humor (N = 9). All articles were assigned to these themes, with some articles providing findings on more than one theme (and thus being counted twice, as well as being listed separately in each theme related table in the appendix).

Habitual Differences in Humor-Related Traits

When looking at differences in humor related traits and "the sense of humor", the best replicated finding is the difference between men and women in self-ratings of their *aggressive* humor style, with men scoring consistently higher (see Cann & Matson, 2014; Cann et al.,

2014; Hall, 2011; Hugelshofer et al., 2006; McCosker & Moran, 2012; Ruch & Heintz, 2016; see Table 1 in the Appendix). Similarly, men obtain higher scores on katagelasticism (the joy of laughing at other) than women (Ruch & Proyer, 2010), with katagelasticism sharing conceptual similarities to aggressive humor style (see Ruch, Esser, Proyer, & Mitrache, 2011). Linking to this, Hehl and Ruch (1985) found that only for males, dominance and aggression predicted the enjoyment of jokes with sexual contents (with males usually finding sexual humor funnier). With respect to the remaining humor styles (Martin et al., 2003) and remaining dispositions towards ridicule and laughter (Ruch & Proyer, 2009), no gender differences were reported. Studies assessing the temperamental basis of the sense of humor, namely trait and state cheerfulness, seriousness, and bad mood, did not find gender differences in the three traits and states (see Ruch et al., 1996; 1997).

This finding – the majority of humor styles and humor-related traits not differing for men and women – is in opposition to the findings of the study by Svebak and colleagues (2004) who found that their "overall humor scores" were higher for men than they were for women. Yet, it should be noted that the latter finding might be due to extremely large sample (increasing the likelihood of a significant result) and the assessment of "overall humor" with a poor method (only three items, with each representing one of the dimensions of the Situational Humor Questionnaire). Studying personality and humor of comedians – a known group of successful humor communicators - Schwehm, McDermut and Thorpe (2015) found that female comedians had higher scores in neuroticism but generally people could not predict whether someone is a comedian or not by the gender of a person, especially as both genders in the comedian sample did not differ in their use of the four humor styles.

Three studies by Hall (2010, 2013, 2015) also investigated the role of humor-related traits in relationships. The first looked at partner embarrassment (Hall, 2010), the second at humor in long-term relationships (Hall, 2013) and the third at humor during the courtship process (Hall, 2015). The first study discovered that the aggressive humor style of a partner

led to embarrassment in the other. Similarly, embarrassment occurred for women whose male partners used self-defeating humor. The second study, also utilizing a measure of the humor styles by Martin et al. (2003), explored "relationship-specific" functions of humor and relationship satisfaction. Here, the "enjoyment function" of humor was shown to mediate the relationship between the two a priori positive humor styles (affiliative, self-enhancing) and satisfaction. The third study used Facebook profiles (Hall, 2015), and face-to-face interactions between opposite sex strangers to see the role humor plays in sexual selection. Humor production, assessed on the Facebook profiles, was linked to extraversion. Successful dating interest in females was associated with male humor production and female's simultaneous laughter. This laughter was rated by trained coders but the paper does not describe how judges evaluated the simultaneous element of the laughter, as opposed to contrived or up-regulated laughter, for example.

All remaining studies focused on the relationship of humor, stress, and stress coping (as well as coping with stress through humor, which was assigned trait-like qualities; see Martin, 1996, for an overview). Firstly, correlational studies looked at the relation of humor styles, coping humor and self-reported stress. In Overholser's study (1992), the Coping Humor Scale (CHS) related positively to lower levels of depression, loneliness, and higher levels of self-esteem. However, retesting after seven weeks showed that these effects were temporary – or that the methodological challenges associated with the CHS did not allow for a reliable testing of long-term effects. The role of humor styles in relation to dysphoria and attributional styles was studied by Hugelshofer and associates (2006). Here, for both genders, fewer depressive symptoms were associated with reporting more affiliative and self-enhancing and less self-defeating humor. Women reporting higher rates of depressed mood also related to the use of aggressive humor (while no such effect was found for men). McCosker and Moran (2012) showed in their study of self-esteem and interpersonal competencies that like with other studies, it was the men who used aggressive humor styles

but there were no differences in any other humor measure. Apart from this, higher self-esteem was predicted by affiliative, self-enhancing, and aggressive humor, as well as lower use of self-defeating humor, for both genders equally. Cann, Watson and Bridgewater (2014) reported that males also use more aggressive humor in the workplace, compared to women. Also investigating the social aspects of humor, Cann and Matson (2014) looked at how the sense of humor, and humor styles are perceived in relation to social desirability. They found that only the sharing of adaptive humor styles is socially desirable and others, like aggressive humor, can be socially damaging (for both genders equally). Utilizing the Multidimensional Sense of Humor Scale of Thorson and Powell (1993), Abel (2002) researched the relationships between humor, coping strategies and stress by selecting people with high and low "sense of humor". In this study, neither humor nor gender related to eight coping strategies or to how individuals dealt with everyday problems. Ford and colleagues (2004) utilized the CHS to investigate stereotype threat on math performance. They found that those women with low coping humor perceived their performance on math tests to be lower than both men or women high in coping humor. Ford and colleagues (2004) further found that state anxiety mediated the relation between stereotype threat and coping sense of humor and the performance a math test.

Another set of studies linked humor-related traits to experimentally induced stress. Lefcourt and colleagues (1997) used the CHS in their study of humor as a moderator of stress level, measured by participant's blood pressure whilst completing stressful tasks. Women higher in coping humor behaviors showed lower systolic blood pressure than those with low coping humor, the reverse was found for men. Abel and Maxwell (2004) investigated gender differences in relation to trait humor and mirth elicitation in tasks designed to be either high or low stress inducing. Watching a humor video was more beneficial for women in low stress but this was opposite for males who profit from the humor conditions under high stress.

Differences in Humor Appreciation

The topic of gender differences in humor appreciation is the most researched theme that emerged. The majority of these studies presents male and female participants with a range of pre-selected humorous stimuli in the lab asking participants to rate the materials along chosen dimensions, as can be seen in Appendix Table 2 (all studies but one: van Giffen & Mahler, 1995, had participants re-telling funny events from their past and had participants rate those stimuli). Variations occur mainly for the mode, structure, and content of the humorous stimuli, the obtained rating dimensions, the inclusion of moderators and mediators, as well as manipulations of the context. To summarize, jokes or cartoons were used most often. Sexual and hostile contents were also predominant, but nonsense and "neutral" humor were also investigated, see Appendix Table 2. Often, the jokes were systematically varied along dimensions, such as explicitness and vulgarity for sexual jokes or cruelty for hostile jokes (e.g., Herzog, 1999; Herzog & Anderson, 2002; Herzog et al., 2006), as well as the victim/target gender in the jokes. Most studies employed ratings of funniness (some typicality, aversiveness, preference, originality) and additionally, the impact of masculinity/femininity, sense of humor, gender identification, sexism, sexual liberation on the result was investigated. Out of eight studies, seven reported that men appreciated sexual humor more than females, sometimes in jokes with mixed targets, in two studies only when the targets were female, see Table 2. Herzog (1999) showed that women appreciated sexual humor with male victims as much as males, while appreciating jokes with females targets less. With respect to hostile humor, consistently, opposite gender target jokes were appreciated more than same gender target jokes, yet while these results were always significant for women, it sometimes failed to be significant for men (e.g., Abrams & Bippus, 2011). Only women showed a preference for affiliative humor instances over hostile or aggressive humor instances, while men rated both types of instances equally positively. No differences in the appreciation of nonsense humor (Köhler & Ruch, 1996; Ruch, 1988),

neutral jokes (Ferstl et al., 2016), or real-life retold funny stories (in which hostile and sexual contents were virtually absent) were found.

Differences in Humor Production and Comprehension Ability

It has been hypothesized that gender differences in humor production may have been shaped through sexual selection, such that humor may signal good quality in genes in prospective mates. Furthermore, it was proposed that men produce humor to enhance mating success (Greengross & Miller, 2011; Tornquist & Chiappe, 2015). Yet, only five studies investigated actual differences in humor production ability, as opposed to self-reported humor production (see Bressler et al., 2006; Tornquist & Chiappe, 2015), see Table 3 in the Appendix. Greengross and Miller (2011) reported that the funniest captions to three cartoons (rated by 6 raters, 4 male and 2 female) produced by men, were on average rated funnier than the captions of women (a finding also reported by Howrigan & MacDonald, 2008 and Mickes et al., 2012) and men produced *more* captions than women. However, in a similar humor production task, Kellner and Benedek (2017) found no gender differences in humor production ability when considering intelligence and creativity (see Table 3). Lastly, Hooper et al. (2016) found no gender differences in two samples with women even being rated funnier than men in a third sample.

Interestingly, all studies that had participants producing humor in the lab aggregated produced humor across different contents for the main analyses. While two studies (Howrigan & MacDonald, 2008; Greengross & Miller, 2011) did not provide information on the contents of the produced humor at all, one study found no differences in humor production across three categories of humorous stimulus material (incongruity-resolution, nonsense, sexual humor) and one further study reported minor differences in the usage of sexual humor and profanity in produced humor (with men using those categories slightly more often than women), but no differences for sexual or profanity contents with respect to the funniness of these captions or the preference by male or female raters occurred (Mickes et al., 2012).

One study investigated humor comprehension: Ferstl, Israel and Putzar (2016) found that men showed different eye movements when reading jokes compared to non-humorous texts, while no such effects occurred for women. In this study, care was taken that the humorous texts did not involve gender stereotypes, sexual or offensive contents and the number of female and male protagonists was equal. Thus, no differences due to the contents of the humor should have occurred.

Differences in Neuroanatomical Responses

When investigating the neural correlates of humor, distinctions should be drawn between the perception and processing of a humorous stimulus, the felt emotion and the behavioral (motor) responses like smiling and laughter (e.g., Goel & Dolan, 2001; Rodden, 2018; Wild et al., 2003). In studies on neuroanatomical processing of humor, responses to different types of stimuli are typically compared, as well as differential responses of men compared to women (see Appendix Table 4). Looking at humor appreciation, two studies utilizing fMRI compared neural responses to funny versus unfunny stimuli and gender differences (Azim et al., 2005; Kohn et al., 2011). Azim and colleagues (2005) found greater left prefrontal cortex (PFC) activation for women when exposed to humorous stimuli compared to non-humorous baselines, but did not find this difference in male participants. Kohn et al. (2011) found that both genders showed stronger activation in the PFC when exposed to humorous stimuli compared to non-humorous stimuli (not just in females, as Azim et al., 2005 reported). Yet, Kohn et al. (2011) found different gender differences to Azim and colleagues: Women showed greater activity in areas involved with limbic reactivity (i.e., appraisal of emotional features of stimuli; amygdala, insula, and anterior cingulate cortex), while men showed greater activation for areas related to the evaluative and executive resources to humor processing (thalamus, dorsolateral prefrontal cortex).

Beyond the comparison of funny versus unfunny stimuli, Chan (2016) investigated neural correlates of sex differences in joke processing considering three types of verbal jokes

(bridge-inference jokes, exaggeration jokes, and ambiguity jokes). The results showed that women exhibited greater cerebral activity in the temporo-parietal-mesiocortical-motor cortex than men while processing the jokes in general. Also, women showed greater activity in the fronto-mesolimbic network while processing "exaggeration jokes" compared to men. "Ambiguity jokes" elicited greater cerebral activity in the frontal paralimbic network in men as compared to women. All joke types elicited greater activation in the anterior prefrontal gyrus of women than in those of men, whereas men showed greater activation than women in the dorsal prefrontal cortex (Chan, 2016; cf. Table 4 in the Appendix).

Differences in Humor Use and Communication

Gender differences in humor use and communication were investigated in a variety of settings (see Table 5 in the Appendix). In relationships, Treger, Sprecher, and Erber (2013) found no gender differences in self-reported humor use in interactions with romantic partners. However, Winterheld, Simpson, and Orina (2013) found during videotaped sessions of relationship conflict, men used two of the humor styles, namely aggressive and affiliative, more often than women (while self-enhancing humor was not coded at all and self-defeating humor was rather coded as "laughing at oneself" *without* a focus on the detrimental consequences of self-deprecating too much and feeling bad whilst doing so). Moreover, women laughed more and were rated less angry when men used affiliative humor (Winterheld et al., 2013). Yet, when coding humor statements in same sex interactions, Dunbar and colleagues (2012) found that women uttered more humorous statements than men. The differences could be attributed to the nature of the dyad (same sex versus mixed) and thus, the results capture gender roles, rather than natural differences in the use of humor.

Women self-reported to use more "positive humor" and "cohesive-building humor" while men reported to use more "aggressive humor", "negative humor", and "outgroup humor", with all small correlations and small effects (Cann, Waterson, & Bridgewater, 2014; Dunbar, Banas, Rodriguez, Liu, & Abra, 2012; Robinson & Smith-Lovin, 2011). Indeed,

certain social pressures may exist for women, such that women were perceived more angrily when using aggressive humor and men responded with laughter when women used more self-defeating humor (Winterheld et al., 2013). This association may be further differentiated depending on status of the female, as males gave greater ratings for long-term attractiveness to high-status female presenters who used self-deprecating humor, but not low-status female presenters (Greengross & Miller, 2008).

In occupational settings, humor communication may be associated with beneficial outcomes for women compared to men. Both helpfulness and humor predicted course ratings for female instructors, while only helpfulness was associated with course ratings in males (Van Giffen, 1990). Moreover, female instructors who used humor during their university lessons obtained higher rates on teaching effectiveness than males (van Giffen, 1990). Furthermore, Williams and Emich (2014) found that females, high in trait-affective perspective taking, exhibited a greater humor self-efficacy after a successful humor attempt. While males high in this trait reported a worse perception of their own humor abilities after failing humor. The authors also evidenced that men felt guiltier after failing a humor attempt than women. Sala, Krupat, and Roter (2002) found a strong positive association between humor and satisfaction in the patient-physician relationship, and the patients of female physicians used more humor than the patients of male physician.

Differences in Responses to Humor

In total, eight studies investigating gender differences in humor responses were included with the majority reporting gender differences (see Appendix Table 6). Only one of them (Martin & Gray, 1996) found that both genders exhibited equivalent responses, such as frequency of laughter or funniness ratings, to humorous stimuli/events. Concerning the exposure to an incomprehensible joke, women showed higher responses of explicit non-understanding, whereas men exhibited more indirect expressions such as questions or interjections (Bell, 2013). Ziv and Gadish (1990) found that younger men instructed to use

humor in a creative task (but not younger women) incorporated more aggressive elements in their written responses than younger men with non-humor instructions. There were also differences in how both genders responded to humor in the context of relationships and interpersonal attraction. Bressler, Martin, and Balshine (2006) more generally asked men and women about the value of humor production and appreciation in a partner and found that while bother genders appreciate a partners' receptivity of their own humor, only women also value a partners humor production ability, a finding that was replicated by Bressler and Balshine (2006), Hone, Hurwitz and Lieberman (2015), as well as by Tornquist and Chiappe (2015) who reported that humor production had stronger effects on women's ratings, in comparison to men's, for partner desirability. Another study, which replicated this pattern, showed that women also reported higher levels of humor appreciation and humor evaluation than men while men reported more production than women (Wilbur & Campbell, 2011). Furthermore, women's evaluations of male's humor are predictive of interest of a romantic relationship (Wilbur & Campbell, 2011). Yet, when looking at the quality of the humorous instance, Cann et al. (2016) claim that only affiliate humor was perceived positively by women, not aggressive humor, while men responded equally to both types of humor. Lastly, Bippus, Dunbar, and Liu (2012) designed different vignettes in which a female or a male used humor to respond to the complaints from their partners. Men rated the vignettes as funnier and more likely to counter-argue than women did (see Appendix Table 6).

Differences in Correlates of Humor

Nine studies investigating gender differences in correlates of humor were included (see Appendix Table 7). The interaction between gender and humor was examined in the context of several psychological constructs, theming around health, interpersonal interactions and the perception of individuals uttering humor.

Concerning the well-researched relationship between humor and health outcomes, some studies have suggested that humor benefits men and women differently. Lefcourt,

Davidson, Prkachin, and Mills (1997) found that certain humor-related dispositions were negatively associated with chronic levels of cardiovascular reactivity among women, whereas it was the opposite (or no relationship) for men. Later, Abel (1998) found that humor moderated the relationship between stress and anxiety, but it occurred only among men. Yet, this correlation only remained significant for men low in humor, while it disappeared for males high in humor. Gender differences were also found when humor was used in health-promoting messages. While men were more persuaded by messages in which humor was combined with high-threatening information, women had more positive attitudes towards those messages with humor and low-threatening information (Hendriks & Janssen, 2017). Finally, no gender benefits/detriments were found in the positive effect of humor on health-related outcomes, such as longevity (Svebak, Romundstad, Holmen, 2010), well-functioning in the work setting (Rupert & Kent, 2007), perceived-stress, frequency of everyday problems, anxiety and the use of coping strategies (Abel, 2002; see Table 7 in the Appendix).

Regarding interpersonal interactions, existing literature has provided mixed evidence about gender differences in humor. For instance, Ziv (1988) exposed that married men and women coincided in indicating the social function of humor as the most relevant to their lives but also agreed on men using more humor in daily lives compared to women. Greengross and Miller (2008) found greater attractiveness ratings to high-status people of the opposite sex when using self-deprecating humor for both genders (see Appendix Table 7).

Discussion

To answer the first research question "(1) In what way has humor been explored and investigated in relation to gender and gender bias through empirical literature? What themes emerge around the research conducted on humor and gender differences?", seven distinct themes emerged from 77 articles meeting the quality criteria. Gender differences can be found in humor-related traits (1), in humor appreciation (2) in humor production and comprehension

ability (3), in neural responses to humor (4), in the use of humor and humor communication (5), in responses to humor (6) and in correlates of humor (7).

Out of the seven themes, five related to differences (or non-existing differences) between women and men that are related mostly to personality and ability (habitual differences, humor appreciation, humor production, humor comprehension, neuroanatomical responses). Three categories, namely "humor use and communication" and "correlates", as well as "humor appreciation" are additionally heavily related to gender roles as well (i.e., sets of behaviors that are assigned to a gender by their reference group or society, see also Kotthoff, 2006) and modulations through the social context. While in the review by Lampert and Ervin-Tipp (1998) few findings were present on the influence of gender roles, gender identification and the influence of the social context, the field has moved forward since then and provided important new insights (e.g., Gray & Ford, 2013; Kochersberger et al., 2014). Also, further studies have appeared that looked into the humor produced in real life situations to establish parallels to the findings from lab studies (Priest & Thein, 2003; van Giffen & Mahler, 1995). Lastly, the often-acclaimed methodological flaws of early works on humor and gender (i.e., use of non-representative materials, de-contextualization, use of "manufactured jokes", see Crawford, 1992, Lampert & Ervin-Tipp, 1998), have been addressed in the recent works. For example, the social context has been included to give the presented humorous materials a frame (such as a facebook profile, with jokes in the status/description, see Strain, Saucier & Martens, 2015) or has been manipulated to strengthen or weaken certain norms (see Cowan & Little, 2013; Gray & Ford, 2013). Also, instead of using jokes studies have asked participants to re-tell funny events from their lives in order to obtain more ecologically valid humorous stimuli (e.g., Abel & Flick, 2012).

Pertaining to differences in humor-related traits, men and women consistently differ in the humor-related traits associated with aggression or aggressive behavior (aggressive humor style, katagelasticism). This is in line with the findings reported by Lampert and Ervin-Tipp

back in 1998 and Kotthoff (2006) and can now be treated as stable and generalizable, as many of the methodological flaws discussed before had been addressed in more recent studies. Moreover, no other differences in humor-related traits were found and thus, men and women describe their sense of humor similarly, with the exception of use of aggressive humor and related traits. While this theme is rather broad, future replications of this systematic reviews may consider splitting this theme into sub-themes, for example, on sub-theme relating to humor styles, one relating to differences in traits associated with the sense of humor, and one sub-theme relating to the impact of humor-related traits to relational uses of humor.

As for humor appreciation, the majority of the studies reported that men appreciated sexual humor more than females (even in mixed gender targets). Both genders appreciated hostile humor with opposite sex targets better, but in some studies, this effect was only significant for women, while the effect were non-significant for men (see Abrams & Bippus, 2011). Thus, joke appreciation in "sensitive topics" (such as sex, hostility) are dependent on the gender of the target and perceiver, with opposite gender targets generally being evaluated more funny. Yet, men seem to have a tendency to be generally less sensitive to target gender, as some studies showed non-significant results for males' ratings of same and opposite gender target jokes with sexual and hostile contents with respect to funniness or less aversiveness towards same gender stimuli as opposed to opposite gender stimuli (or stronger effects in female subsamples; see Table 2). No differences in nonsense humor were found.

To conclude, the theme "sexuality" identified by Kotthoff (2006) is still emergent and thus stable. In lab studies with pre-selected stimuli, gender differences in humor appreciation are apparent (even after ruling out serious methodological flaws of very early studies) and mostly consistent. Women seem to appreciate emotionally arousing stimuli (hostile, sexual humor) less than men, while no differences were found to joke structures. In re-told funny events, hostile and sexual contents were extremely seldom and not surprisingly, no differences between men and women are reported. The later observation led to the question,

whether these results obtained in the lab could be accounted to effects of social desirability (see Aillaud & Piolat, 2012). Aillaud and Piolat (2012) argued that men and women may not really differ in their perception of humorous materials, but rather merely behave in accordance with the cultural beliefs and roles they consider important– we would add – *in that moment*. If participants try to behave in accordance with what they think is expected of them, cultural beliefs and norms may be activated and influence the participants' behavior. Consequently, if it is less clear to the participant what the experiment is about, or what the beliefs and roles are that the experimenter tries to get, or gender irrelevant norms are wanted, gender differences should be smaller. Although this remains to be tested, the finding by Gray and Ford (2013) may point in this direction. In their study, the social context (the local norms) and the joke type where manipulated, and the results show that when men and women refer to a norm that is not explicitly related to gender, or when participants do not build the hypothesis that they should behave according to their gender roles, the gender differences disappear.

Only very few studies actually assessed humor production as in generating funny punch-lines and for those studies, the same amount of studies reported gender differences as studies that did not find gender differences (see Greengross & Miller, 2011; Hooper et al., 2016; Kellner & Benedek, 2017). Interestingly, judging the produced humor of males and females also led to different results when comparing different nation (but same language) groups of raters. Thus, if men and women are found to be unequally funny, this effect could be specific to a cultural group of raters and – as the study of Hooper et al. (2016) showed, might not even be replicable in comparable language groups (i.e., different nations of English speakers). Yet, the sample in the Mickes et al. (2012) study, as well as all three samples in the Hooper et al. (2016) study agreed on generally *thinking* that men were funnier than women in holistic self-ratings, as well as when having to memorize the author gender of the funniest versus least funny joke. Thus, men and women may think that men are funnier (the attribution bias was replicated in different studies) but whether they actually are funnier is to be debated.

All studies on differences in the neural processing reported gender differences and found differential effects for prefrontal cortex activations (i.e., for different prefrontal cortex regions). Thus, while the studies agree that gender differences in neural responses exist, there is no agreement yet on the regions the differences occur, possibly also due to the fact that the studies used different types of stimuli, different sets of content types and different types of comparison stimuli (i.e., baseline comparison, comparison to non-humorous stimuli or other types of humorous stimuli).

For humor use and communication, one needs to be careful when looking at the existing studies, as many studies rely on self-reported use and communication of humor, and not actual observations. The methodologically more convincing studies using behavioral coding and observations (e.g. Dunbar, Banas, Rodriguez, Liu, & Abra, 2012) suggest gender differences, yet, they are not unidirectional. While in mixed sex dyads, men utter humor more frequently than women. In same sex dyads, women utter more humor than men. Thus, there are no general gender differences, but the differences evolve from different social context and may thus be more accountable to gender roles rather than natural differences in humor use between men and women. It is just not as simple as "men use more humor than women" as it depends on who the individuals are talking to and in which situation they are in (romance, workplace). While findings that men produce more humor than females in romantic relationships and courtship settings has been replicated in several studies, it does not generalize across all situations. Indeed, several theoretical accounts try to explain why men produce more humor in mixed sex dyads (i.e., humor being a proxy for intelligence and thus fitness, potentially leading to more mating success, see for example Grammer & Eibl-Eibesfeldt, 1990, Bressler et al., 2006, Greengross & Miller, 2011; Tornquist & Chiappe, 2015). Yet, women telling funny stories (and thus producing) humor also enhanced their attractiveness in an experimental design, although self-report data suggests that humor production ability is much more important for mating success in males. While currently, this

theme is thematically broad, with a growing body of literature, this theme might be split into sub-themes in future replications of the systematic review.

With respect to responses to humor, findings consistently showed that while both genders highly value a sense of humor in partners and appreciate a partner's receptivity of one's own humor (Bressler et al. 2006; Tornquist & Chiappe, 2015), only women valued a partner's humor production ability, and rated men more attractive when they produced humor in a courtship attempt. Thus, at least in courtship situations, traditional gender roles prevail when it comes to humor. Yet, Gueguen (2010) showed that women did not perceive the men any more attractive in the humor courtship condition of the study, which highlighted how humor use was mediated by perceived funniness and sociability. So, the compliance of the female in the humor condition could be a social compliance gesture rather than a straightforward indicator of successful courtship solicitation.

While findings on correlates of humor are mixed, some common denominators can be extracted: Firstly, humor is valued as important in interactions and serves different social functions. Secondly, if gender differences were found, they sometimes only occurred in individuals low in the sense of humor. Thus, gender differences might only be apparent in low sense of humor. Third, gender roles (e.g., women appreciating humor, men producing humor) explain some of the findings on humor correlates: for example, men judge their own failed attempt in humor as worse than women (see Williams & Emich, 2014).

Next, the second and third research questions: (2) "What is the quality of the evidence around humor and gender differences?" and (3) "Is there a role in research methodology on the findings of gender differences?" are answered. To address question 2, all articles included in this review obtained a QUALSYST score of > 0.75. This score may be considered as an indicator of good quality, but also a more conservative cut-off (e.g., Kmet et al., 2011). Of the articles that were eligible for the QUALSYST rating 49% reached this cut-off. Future studies

and updates of a review on humor and gender may apply a more liberal cut-off to see whether the emerging themes can be replicated or need to be broadened.

With regards to question 3 ("Is there a role in research methodology on the findings of gender differences?"), several aspects can be looked at. First, the applied research methods and designs systematically vary across the themes. To assess habitual differences in humor-related traits, humor styles and the sense of humor (theme 1), all studies utilized self-report measures and most studies applied a cross-sectional design. Thus, 100% of the studies in theme one utilized this method. Yet, some studies applied tests and behavioral methods to assess further outcomes (e.g., math performance, pain threshold). For gender differences in humor appreciation, humor production and comprehension, as well as neuroanatomical differences (themes 2, 3, 4) all studies applied objective methods, such as tests and neuroimaging techniques, sometimes coupled with self-report data on humor-related traits. The themes humor use and communication, differences in responses to humor and humor correlates (themes 5, 6, 7) were more heterogeneous with respect to the utilized methods and designs. In many cases, studies relied on self-report data, for example on the self-reported use of humor or humor as a communication tool. In these themes, future studies should aim at providing hard data to complement, replicate or extend the existing findings (i.e., utilizing ambulatory assessment techniques to actually count instances of humor in daily interactions and conversations and to be able to analyze the quality of those utterances). Second, another interesting aspect would be whether certain methods systematically go along with finding gender differences (or not finding them), across themes, while others do not. For example, would self-reports systematically report gender differences while other methods – for example objective methods or tests - would not? Luckily, this notion could not be supported by our findings, although self-reports on the perception of the funniness of men and women tend to accentuate more prominent differences between males and females, in accordance with gender stereotypes and roles (in which men tend to be seen as funnier than women).

Limitations

Overall, the current systematic review is not comprehensive, as it follows a standardized procedure guided by a priori research questions. For this reason, some papers that potentially are informative on humor and gender differences might not be represented here. In more details, firstly, in the initial literature search strategy, we chose to narrow the search by requiring both terms (humor and gender) to be present (by linking them with the Boolean operator "and"), as we wanted to focus on papers that were looking at gender differences as the primary focus. This was also clearly stated in our eligibility criteria and further checked after the initial search was carried out. Yet, as a side effect of this strategy, we might have missed some papers that have investigated aspects of humor and did report gender differences in the paper, but only as a minor outcome of the research. Secondly, while we did not specifically exclude papers with qualitative data in our search, our search criteria might have led to a bias against qualitative data as none of the authors is specialized in the analysis of qualitative data and all of the authors are psychologists, in line with the limitations already pointed out by Kmet et al. (2014). Thus, we could have potentially been biased against qualitative methodologies, such as discourse analysis or conversation analysis, even though the criteria were applied that Kmet et al. suggested. Thirdly, we did not include studies that reported data only on males or only on females, as such studies do not allow for a direct comparison of men and women. Yet, such studies might still be informative on the humor of men and women which could be compared as part of a comprehensive review.

Conclusion

Small but replicable gender differences exist in some areas of humor, the fore front humor appreciation and responses to humor and humor communication. Yet, the size of these findings may be heavily influenced by the method of assessment and the context (in line with Lampert & Ervin-Tipp, 1998). Especially the role of gender roles on self-reports and norm

pressure in lab studies needs to be explored in more detail in future studies (see also Ergül, 2014), as well as the role of the gender of raters of humor (e.g., judges of funniness in humor production studies) on the results. Moreover, while some gender differences exist and can be replicated across studies and language groups, still little is known on gender differences in humor in a) natural settings, b) cultures outside the Western world, and in the c) development of humor across the life span, another endeavor ready to be pursued in future studies.

Ethical Approval

No data collection was conducted in this article and therefore, no ethic approval was necessary.

Informed Consent

No data collection was conducted in this article and thus no informed consent was necessary.

Conflict of Interest Statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References (including Appendix)

Abel, M. H. (1998). Interaction of humor and gender in moderating relationships between stress and outcomes. *The Journal of Psychology: Interdisciplinary and Applied, 132*(3), 267-276. doi:10.1080/00223989809599166

Abel, M. H. (2002). Humor, stress, and coping strategies. *Humor: International Journal of Humor Research, 15*(4), 365-381. doi:10.1515/humr.15.4.365

Abel, M. H., & Flick, J. (2012). Mediation and moderation in ratings of hostile jokes by men and women. *Humor: International Journal of Humor Research, 25*(1), 41-58. doi:10.1515/humor-2012-0003

Abel, M. H., & Maxwell, D. (2002). Humor and affective consequences of a stressful task. *Journal of Social and Clinical Psychology, 21*(2), 165-190. doi:10.1521/jscp.21.2.165.22516

Abrams, J. R., & Bippus, A. M. (2011). An intergroup investigation of disparaging humor. *Journal of Language and Social Psychology, 30*(2), 193-201.

Abrams, J. R., & Bippus, A. M. (2014). Gendering jokes: Intergroup bias in reactions to same- versus opposite-gender humor. *Journal of Language and Social Psychology,*

33(6), 692-702. doi:10.1177/0261927X14544963

Abrams, J. R., Bippus, A. M., & McGaughey, K. J. (2015). Gender disparaging jokes: An investigation of sexist-nonstereotypical jokes on funniness, typicality, and the moderating role of ingroup identification. *Humor: International Journal of Humor Research, 28*(2), 311-326. doi:10.1515/humor-2015-0019

Aillaud, M., & Piolat, A. (2012). Influence of gender on judgment of dark and nondark humor. *Individual Differences Research, 10*(4), 211-222.

American Psychological Association. 2009. *Publication manual* (sixth Edition). Washington. DC: American Psychological Association.

Azim, E., Mobbs, D., Jo, B., Menon, V., & Reiss, A. L. (2005). Sex differences in brain activation elicited by humor. *Proceedings of the National Academy of Sciences of the United States of America, 102*(45), 16496-16501.

Bell, N. D. (2013). Responses to incomprehensible humor. *Journal of Pragmatics, 57*, 176-189. doi:10.1016/j.pragma.2013.08.019

Bippus, A. M., Dunbar, N. E., & Liu, S.-J. (2012). Humorous responses to interpersonal complaints: Effects of humor style and nonverbal expression. *The Journal of Psychology: Interdisciplinary and Applied*, *146*(4), 437-453. doi:10.1080/00223980.2011.652696

Bressler, E. R., & Balshine, S. (2006). The influence of humor on desirability. *Evolution and Human Behavior, 27*(1), 29-39. doi:10.1016/j.evolhumbehav.2005.06.002

Bressler, E. R., Martin, R. A., & Balshine, S. (2006). Production and appreciation of humor as sexually selected traits. *Evolution and Human Behavior, 27*(2), 121-130. doi:10.1016/j.evolhumbehav.2005.09.001

Cann, A., & Matson, C. (2014). Sense of humor and social desirability: Understanding how humor styles are perceived. *Personality and Individual Differences, 66*, 176-180. doi:10.1016/j.paid.2014.03.029

Cann, A., Cann, A. T., & Jordan, J. A. (2016). Understanding the effects of exposure to humor expressing affiliative and aggressive motivations. *Motivation and Emotion, 40*(2), 258-267. doi:10.1007/s11031-015-9524-8

Cann, A., Watson, A. J., & Bridgewater, E. A. (2014). Assessing humor at work: The Humor Climate Questionnaire. *Humor: International Journal of Humor Research, 27*(2), 307-323. doi:10.1515/humor-2014-0019

Carretero-Dios, H., & Ruch, W. (2010). Humor appreciation and sensation seeking: Invariance of findings across culture and assessment instrument? *Humor: International Journal of Humor Research, 23*(4), 427-445. doi:10.1515/HUMR.2010.020

Chan, Y. C. (2016). Neural correlates of sex/gender differences in humor processing for different joke types. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2015.00536

Conway, M., & Dube, L. (2002). Humor in persuasion on threatening topics: Effectiveness is a function of audience sex role orientation. *Personality and Social Psychology Bulletin, 28*(7), 863-873. doi:10.1177/01467202028007001

Cowan, M. L., & Little, A. C. (2013). The effects of relationship context and modality on ratings of funniness. *Personality and Individual Differences, 54*(4), 496-500. doi:10.1016/j.paid.2012.10.020

Crawford, M. (1992). Just kidding: Gender and conversational humor. In R. Barreca (Ed.), *New perspectives on women and comedy* (pp. 23-38). Philadelphia: Gondon and Breach.

Dunbar, N. E., Banas, J. A., Rodriguez, D., Liu, S.-j., & Abra, G. (2012). Humor use in power-differentiated interactions. *Humor: International Journal of Humor Research, 25*(4), 469-489. doi:10.1515/humor-2012-0025

Ergül, H. (2014). Gender roles in humor. In S. Attardo (Ed.), *Encyclopedia of humor studies* (pp. 262-263). Thousand Oaks, CA: Sage.

Ferstl, E. C., Israel, L., & Putzar, L. (2017). Humor facilitates text comprehension: Evidence from eye movements. *Discourse Processes, 54*(4), 259-284. doi:10.1080/0163853X.2015.1131583

Ford, T. E., Ferguson, M. A., Brooks, J. L., & Hagadone, K. M. (2004). Coping Sense of Humor Reduces Effects of Stereotype Threat on Women's Math Performance. *Personality and Social Psychology Bulletin, 30*(5), 643-653. doi:10.1177/0146167203262851

Goel, V., & Dolan, R.J. (2001). The functional anatomy of humor: Segregating cognitive and affective components. *Nature Neuroscience, 4*, 237–238.

Grammer, K., & Eibl-Eibesfeldt, I. (1990). The ritualisation of laughter. *Natürlichkeit der Sprache und der Kultur, 18,* 192-214.

Gray, J. A., & Ford, T. E. (2013). The role of social context in the interpretation of sexist humor. *Humor: International Journal of Humor Research, 26*(2), 277-293. doi:10.1515/humor-2013-0017

Greengross, G., & Miller, G. (2011). Humor ability reveals intelligence, predicts mating success, and is higher in males. *Intelligence, 39*(4), 188-192. doi:10.1016/j.intell.2011.03.006

Greengross, G., & Miller, G. F. (2008). Dissing oneself versus dissing rivals: Effects of status, personality, and sex on the short-term and long-term attractiveness of self-deprecating and other-deprecating humor. *Evolutionary Psychology, 6*(3), 393-408.

Greenwood, D., & Isbell, L. M. (2002). Ambivalent sexism and the dumb blonde: Men's and women's reactions to sexist jokes. *Psychology of Women Quarterly, 26*(4), 341-350. doi:10.1111/1471-6402.t01-2-00073

Guéguen, N. (2010). Men's sense of humor and women's responses to courtship solicitations: An experimental field study. *Psychological Reports, 107*(1), 145-156. doi:10.2466/07.17.PR0.107.4.145-156

Hall, J. A. (2011). Is it something I said? Sense of humor and partner embarrassment. Journal of Social and Personal Relationships, 28(3), 383-405. doi:10.1177/0265407510384422

Hall, J. A. (2013). Humor in long-term romantic relationships: The association of general humor styles and relationship-specific functions with relationship satisfaction. *Western Journal of Communication, 77*(3), 272-292. doi:10.1080/10570314.2012.757796

Hall, J. A. (2015). Sexual selection and humor in courtship: A case for warmth and extroversion. *Evolutionary Psychology, 13*(3), 1-10. doi:10.1177/1474704915598918

Hehl, F.-J., & Ruch, W. (1985). The location of sense of humor within comprehensive personality spaces: An exploratory study. *Personality and Individual Differences, 6*(6), 703-715. doi:10.1016/0191-8869(85)90081-9

Hendriks, H., & Janssen, L. (2017). Frightfully funny: combining threat and humour in health messages for men and women. *Psychology & Health*, 1-20. doi:10.1080/08870446.2017.1380812

Herzog, T. R. (1999). Gender differences in humor appreciation revisited. *Humor: International Journal of Humor Research, 12*(4), 411-423. doi:10.1515/humr.1999.12.4.411

Herzog, T. R., & Anderson, M. R. (2000). Joke cruelty, emotional responsiveness, and joke appreciation. *Humor: International Journal of Humor Research, 13*(3), 333-351. doi:10.1515/humr.2000.13.3.333

Herzog, T. R., & Hager, A. J. (1995). The prediction of preference for sexual cartoons. *Humor: International Journal of Humor Research, 8*(4), 385-405. doi:10.1515/humr.1995.8.4.385

Herzog, T. R., Harris, A. C., Kropscott, L. S., & Fuller, K. L. (2006). Joke cruelty and joke appreciation revisited. *Humor: International Journal of Humor Research, 19*(2), 139-156. doi:10.1515/HUMOR.2006.007

Hofmann, J., & Ruch, W. (2017). Laughter. In D. S. Dunn (Ed.), *Oxford Bibliographies Online*. New York: Oxford University Press. www.oxfordbibliographies.com

Hone, L. S. E., Hurwitz, W., & Lieberman, D. (2015). Sex differences in preferences for humor: A replication, modification, and extension. *Evolutionary Psychology, 13*(1), 167-181. doi:10.1177/147470491501300110

Hooper, J., Sharpe, D., & Roberts, S. G. B. (2016). Are men funnier than women, or do we just think they are? *Translational Issues in Psychological Science, 2*(1), 54-62. doi:10.1037/tps0000064

Howrigan, D. P., & MacDonald, K. B. (2008). Humor as a mental fitness indicator. *Evolutionary Psychology, 6*(4), 625-666. doi:10.1177/147470490800600411

Hugelshofer, D. S., Kwon, P., Reff, R. C., & Olson, M. L. (2006). Humour's role in the relation between attributional Style and dysphoria. *European Journal of Personality, 20*(4), 325-336. doi:10.1002/per.586

Kellner, R., & Benedek, M. (2017). The role of creative potential and intelligence for humor production. *Psychology of Aesthetics, Creativity, and the Arts, 11*(1), 52-58. doi:10.1037/aca0000065

Khan, K, Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, 96(3), 118-121.

Kmet, L. M., Cook, L. S. & Lee. R. C. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields* (Vol. 22). Edmonton: Alberta Heritage Foundation for Medical Research.

Kochersberger, A. O., Ford, T. E., Woodzicka, J. A., Romero-Sanchez, M., & Carretero-Dios, H. (2014). The role of identification with women as a determinant of amusement with sexist humor. *Humor: International Journal of Humor Research, 27*(3), 441-460. doi:10.1515/humor-2014-0071

Kohler, G., & Ruch, W. (1996). Sources of variance in current sense of humor inventories:

How much substance, how much method variance? *Humor: International Journal of Humor Research, 9*(3-4), 363-397. doi:10.1515/humr.1996.9.3-4.363

Kohn, N., Kellermann, T., Gur, R. C., Schneider, F., & Habel, U. (2011). Gender differences in the neural correlates of humor processing: Implications for different processing modes. *Neuropsychologia, 49*(5), 888-897. doi:10.1016/j.neuropsychologia.2011.02.010

Kotthoff, H. (2006). Gender and humor: The state of the art. *Journal of Pragmatics, 38*, 4-25.

Lampert, M. & Ervin-Tipp, S. (1998). Exploring paradigms: the study of gender and sense of humor near the end of the 20th century. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 231-270). Berlin, Germany: Mouton de Gruyter.

Lefcourt, H. M., Davidson, K., Prkachin, K. M., & Mills, D. E. (1997). Humor as a stress moderator in the prediction of blood pressure obtained during five stressful tasks. *Journal of Research in Personality, 31*(4), 523-542. doi:10.1006/jrpe.1997.2191

Martin, G. N., & Gray, C. D. (1996). The effects of audience laughter on men's and women's responses to humor. *Journal of Social Psychology, 136*(2), 221-231.

Martin, R. A. (1996). The situational humor response questionnaire (SHRQ) and coping humor scale (CHS): A decade of research findings. *Humor: International Journal of Humor Research* 9, 251-272.

Martin, R. A. (2010). *The psychology of humor: An integrative approach*. Burlington: Elsevier.

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J& Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality, 37*(1), 48-75. doi:10.1016/s0092-6566(02)00534-2

McCosker, B., & Moran, C. C. (2012). Differential effects of self-esteem and interpersonal competence on humor styles. *Psychology Research And Behavior Management, 5,* 143-150. doi:10.2147/PRBM.S36967

Mickes, L., Walker, D. E., Parris, J. L., Mankoff, R., & Christenfeld, N. J. S. (2012). Who's funny: Gender stereotypes, humor production, and memory bias. *Psychonomic Bulletin & Review, 19*(1), 108-112. doi:10.3758/s13423-011-0161-2

Mundorf, N., Bhatia, A., Zillmann, D., Lester, P., & Robertson, S. (1988). Gender differences in humor appreciation. *Humor: International Journal of Humor Research, 1*(3), 231-243. doi:10.1515/humr.1988.1.3.231

Overholser, J. C. (1992). Sense of humor when coping with life stress. *Personality and Individual Differences, 13*(7), 799-804. doi:10.1016/0191-8869(92)90053-r

Priest, R. F., & Thein, M. T. (2003). Humor appreciation in marriage: Spousal similarity, assortative mating, and disaffection. *Humor: International Journal of Humor Research, 16*(1), 63-78. doi:10.1515/humr.2003.005

Proyer, R. T., & Ruch, W. (2010). Enjoying and fearing laughter: Personality characteristics of gelotophobes, gelotophiles, and katagelasticists. *Psychological Test and Assessment Modeling, 52*(2), 148-160.

Robinson, D. T., & Smith-Lovin, L. (2001). Getting a laugh: Gender, status, and humor in task discussions. *Social Forces, 80*(1), 123-158. doi:10.1353/sof.2001.0085

Romundstad, S., Svebak, S., Holen, A., & Holmen, J. (2016). A 15-Year Follow-Up Study of Sense of Humor and Causes of Mortality: The Nord-Trondelag Health Study. *Psychosomatic Medicine, 78*(3), 345-353. doi:10.1097/psy.0000000000000275

Rodden, F. A. (2018). The neurology and psychiatry of humor, smiling, and laughter: a tribute to Paul McGhee Part II. neurological studies and brain imaging. *Humor: International Journal of Humor Research, 31*(2), 373–399.

Ruch, W. (1988). Sensation seeking and the enjoyment of structure and content of humour: Stability of findings across four samples. *Personality and Individual Differences, 9*(5), 861-871. doi:10.1016/0191-8869(88)90004-9

Ruch, W. (1998). Foreword and overview. Sense of humor: A new look at an old concept. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 3-14). Berlin, Germany: Mouton de Gruyter.

Ruch, W. (2008). The psychology of humor. In V. Raskin (Ed.), *The primer in humor research* (pp. 17 - 100). Berlin, Germany: Mouton de Gruyter.

Ruch, W., & Heintz, S. (2016). The German version of the Humor Styles Questionnaire: Psychometric properties and overlap with other styles of humor. *Europe's Journal of Psychology, 12*(3), 434-455. doi:10.5964/ejop.v12i3.1116

Ruch, W. & Hofmann, J. (2013). Humor. In D. S. Dunn (Ed.), *Oxford Bibliographies Online*. New York: Oxford University Presswww.oxfordbibliographies.com. doi:10.1093/obo/9780199828340-0127

Ruch, W., Kohler, G., & vanThriel, C. (1996). Assessing the "humorous temperament": Construction of the facet and standard trait forms of the State-Trait-Cheerfulness-inventory - STCI. *Humor: International Journal of Humor Research, 9*(3-4), 303-339. doi:10.1515/humr.1996.9.3-4.303

Ruch, W., Kohler, G., & vanThriel, C. (1997). To be in good or bad humour: Construction of the state form of the State-Trait-Cheerfulness-Inventory-STCI. *Personality and Individual Differences, 22*(4), 477-491. doi:10.1016/s0191-8869(96)00231-0

Ruch, W., Proyer, R. T., Esser, C., & Mitrache, O. (2011). Cheerfulness and everyday humorous conduct. In Romanian Academy, “George Barit“ Institute of History, Department of Social Research (Ed.), *Studies and researches in social sciences* (pp. 67-87, Vol. 18). Cluj-Napoca: Argonaut Press.

Rupert, P. A., & Kent, J. S. (2007). Gender and work setting differences in career-sustaining behaviors and burnout among professional psychologists. *Professional Psychology-Research and Practice, 38*(1), 88-96. doi:10.1037/0735-7028.38.1.88

Sala, F., Krupat, E., & Roter, D. (2002). Satisfaction and the use of humor by physicians and patients. *Psychology & Health, 17*(3), 269-280. doi:10.1080/08870440290029520

Schwehm, A. J., McDermut, W., & Thorpe, K. (2015). A gender study of personality and humor in comedians. *Humor: International Journal of Humor Research, 28*(3), 427-448. doi:10.1515/humor-2015-0069

Strain, M., Saucier, D., & Martens, A. (2015). Sexist Humor in Facebook Profiles: Perceptions of Humor Targeting Women and Men. *Humor: International Journal of Humor Research, 28*(1), 119-141. doi:10.1515/humor-2014-0137

Svebak, S., Martin, R. A., & Holmen, J. (2004). The prevalence of sense of humor in a large, unselected county population in Norway: Relations with age, sex, and some health indicators. *Humor: International Journal of Humor Research, 17*(1-2), 121-134. doi:10.1515/humr.2004.001

Svebak, S., Romundstad, S., & Holmen, J. (2010). A 7-year prospective study of sense of humor and mortality in an adult county population: the hunt-2 study. *International Journal of Psychiatry in Medicine, 40*(2), 125-146. doi:10.2190/PM.40.2.a

Thorson, J.A., & Powell, F.C. (1993). Development and validation of a multidimensional sense of humor scale. *Journal of Clinical Psychology, 49*, 13–23. doi:10.1002/1097-4679(01)49:1<13::AID-JCLP2270490103>3.0.CO;2-S

Tornquist, M., & Chiappe, D. (2015). Effects of Humor Production, Humor Receptivity, and Physical Attractiveness on Partner Desirability. *Evolutionary Psychology, 13*(4). doi:10.1177/1474704915608744

Treger, S., Sprecher, S., & Erber, R. (2013). Laughing and liking: Exploring the interpersonal effects of humor use in initial social interactions. *European Journal of Social Psychology, 43*(6), 532-543. doi:10.1002/ejsp.1962

Van Giffen, K. (1990). Influence of professor gender and perceived use of humor on course evaluations. *Humor: International Journal of Humor Research, 3*(1), 65-73. doi:10.1515/humr.1990.3.1.65

Van Giffen, K., & Maher, K. M. (1995). Memorable humorous incidents: Gender, themes and setting effects. *Humor: International Journal of Humor Research, 8*(1), 39-50. doi:10.1515/humr.1995.8.1.39

Wilbur, C. J., & Campbell, L. (2011). Humor in romantic contexts: do men participate and women evaluate? *Personality and Social Psychology Bulletin, 37*(7), 918-929. doi:10.1177/0146167211405343

Wild, B., Rodden, F.A., Grodd, W., & Ruch, W. (2003). Neural correlates of laughter and humor. *Brain: A Journal of Neurology, 126*, 2121–2138.

Williams, M., & Emich, K. J. (2014). The experience of failed humor: implications for interpersonal affect regulation. *Journal of Business and Psychology, 29*(4), 651-668. doi:10.1007/s10869-014-9370-9

Winterheld, H. A., Simpson, J. A., & Orina, M. M. (2013). It's in the way that you use it: attachment and the dyadic nature of humor during conflict negotiation in romantic couples. *Personality and Social Psychology Bulletin, 39*(4), 496-508. doi:10.1177/0146167213479133

Ziv, A. (1988). Humor's role in married life. *Humor: International Journal of Humor Research, 1*(3), 223-229. doi:10.1515/humr.1988.1.3.223

Ziv, A., & Gadish, O. (1990). The disinhibiting effects of humor: Aggressive and affective responses. *Humor: International Journal of Humor Research, 3*(3), 247-257. doi:10.1515/humr.1990.3.3.247

Figures

Figure 1. Search Strategy and Outcomes.